

Evaluation of a Tool to Categorize Patients by Reading Literacy and Computer Skill to Facilitate the Computer-Administered Patient Interview

David F. Lobach, MD, PhD, MS¹; Victor Hasselblad, PhD²; Barbara M. Wildemuth, PhD³

¹Division of Clinical Informatics, Department of Community and Family Medicine and

²Department of Biostatistics and Bioinformatics, Duke University, Durham, NC

³School of Information and Library Science, University of North Carolina, Chapel Hill, NC

ABSTRACT

Past efforts to collect clinical information directly from patients using computers have had limited utility because these efforts required users to be literate and facile with the computerized information collecting system. In this paper we describe the creation and use of a computer-based tool designed to assess a user's reading literacy and computer skill for the purpose of adapting the human-computer interface to fit the identified skill levels of the user. The tool is constructed from a regression model based on 4 questions that we identified in a laboratory study to be highly predictive of reading literacy and 2 questions predictive of computer skill. When used in 2 diverse clinical practices the tool categorized low literacy users so that they received appropriate support to enter data through the computer, enabling them to perform as well as high literacy users. Confirmation of the performance of the tool with a validated reading assessment instrument showed a statistically significant difference ($p=0.0025$) between the two levels of reading literacy defined by the tool. Our assessment tool can be administered through a computer in less than two minutes without requiring any special training or expertise making it useful for rapidly determining users' aptitudes.

INTRODUCTION

Computers have been used to collect clinical information directly from patients for over 3 decades.^{1,8} The advantages as well as the disadvantages of the computer-administered patient interview have recently been extensively reviewed.⁹ In general, computers have been shown to collect a larger quantity of historical information and more sensitive information than clinician-administered interviews or paper forms.^{3,7,9} However, almost all of the previous attempts at the computer-administered patient interview required patients to be able to read and to have some familiarity with a computer.^{1,3,8} We have sought to bridge the digital divide for the computer-administered patient interview by creating a data collection system that can be used by all patients, adapting its interface to fit the reading literacy, computer skills and native language of the user.^{10,11} In order to provide appropriate support for

reading and computer skills through the adaptable interface, however, we needed a mechanism to assess the user's reading and computer skills expeditiously. The few systems that have sought to support users with low reading literacy and minimal computer skills¹²⁻¹⁴ supported all users at the same level which could be frustrating and inefficient for users who did not require this level of support. Several paper-based instruments are available to assess reading literacy quickly,¹⁵⁻¹⁸ but no tools explicitly seek to categorize users by reading literacy and computer skill for the purpose of adapting the human-computer interface to fit the skills of the user.¹⁹

The purpose of this study was to develop and evaluate a tool to assess and categorize patients according to their reading literacy (high versus low) and computer skills (high versus low). This study was part of a larger project, known by the acronym MADELINE (Multimedia Adaptive Data Entry and Learning Interface in a Networked Environment), undertaken to collect health information directly from patients using a computer irrespective of the patient's reading ability or computer experience.¹¹

METHODS

Question Identification. With direction from a literacy expert and a computer usability specialist, we created a set of 16 questions which were believed to be highly predictive of reading and computer literacy based on the use of these questions in other survey instruments such as the National Adult Literacy Survey (NALS),²⁰ or on face validity regarding user preferences. Information collected pertaining to reading literacy included age, educational level, income level, employment status, type of occupation, amount of time worked in the past year, number of books read per week, and disabilities. Questions pertaining to computer skill asked for number of hours of computer use per week, comfort level using a mouse, comfort level entering data with a keyboard, and types of activities done on a computer, if any. Subjects were also asked a reading preference question (in response to a sample clinical question, subjects were asked if they would prefer to read the questions themselves or have the questions read to them) and a computer preference question (subjects were asked if they preferred to respond to questions

one at a time using a touch screen or to several questions on the screen using a mouse and keyboard).

Exploratory Evaluation. The test question set was evaluated in an exploratory laboratory study involving 100 subjects. Subjects were recruited through two clinic sites and through community organizations which offered reading instruction to low literacy individuals. Potential participants were screened by telephone with questions to assess their reading ability to ensure an adequate representation of subjects who were likely to have limited reading ability. In the laboratory, all subjects were read the test question set by a research assistant. The subject's reading literacy was then assessed using the TOWRE¹⁶ (Test of Word Reading Efficiency, PRO-ED, Austin, Texas) Sight Word Efficiency Score and the Phonemic Decoding Efficiency Score. Raw scores were converted into a grade equivalent.

Computer literacy was assessed using the Computer Self-Efficacy Scale.^{21,22} For this instrument, subjects were read a series of statements related to their self-efficacy with computers. For each statement they were to indicate the response that best described their current belief using a five-point Likert scale.

Creation of the Assessment Tool. The questions which were most predictive of reading literacy and computer skill were incorporated into a prediction model (below). These questions were also integrated into the introduction module of the MADELINE system and presented to patients with maximal reading and computer support, i.e., one question per screen with audio files to read the questions, and touch screen data entry (Figure 1). By processing patients' responses to the assessment questions through the model, patients were assigned to receive the appropriate level of reading or computer support to enable them to complete a 75 item health maintenance and risk assessment questionnaire.

Confirmatory Evaluation. A convenience sample of subjects enrolled in a separate study to assess the effectiveness of the MADELINE system was selected for inclusion in a sub-study to confirm the validity of the model for predicting reading literacy and computer skill. These 78 subjects underwent a literacy assessment using the WRAT3¹⁵ (Wide Range Achievement Test Version 3, Wide Range, Inc., Wilmington, DE). We used this instrument (instead of TOWRE) as our gold standard for reading literacy for this evaluation because we needed a brief test to use with the already lengthy set of forms and surveys included in the evaluation of MADELINE; because we wanted a different validated literacy assessment standard than we had used to create the model; and because we believed that the WRAT3 might more closely approximate the reading skills that we were seeking to assess with our tool, i.e., the skill required

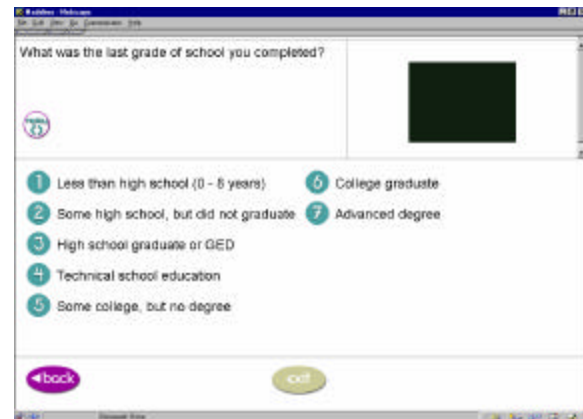


Figure 1. Screen shot from the MADELINE assessment tool.

to complete a questionnaire. The literacy assignment using the model was done using the patients' responses to the assessment questions programmed into the MADELINE system.

Statistical Analysis. The contribution of each assessment question for predicting reading literacy or computer skill was determined through stepwise linear regression using SAS PROC GLM (SAS Institute, Cary, NC). The confirmation of the model's effectiveness for correctly predicting reading literacy was done by comparing differences from the WRAT3 test results using the Wilcoxon rank-sum test.²³

RESULTS

Model Development. Stepwise linear regression analysis was used to estimate a model that would predict reading literacy (as measured by the TOWRE Sight Word Score) from baseline variables. Because the relationship of the Sight Word Score to educational level appeared to be non-linear, we included a separate dummy variable for the first (low) educational level and then assumed that the changes were linear in level for subsequent education categories. Five variables entered the model and these are shown in Table 1.

The model was then used to predict the Sight Word Score for every person. If the predicted score was less than 75, then the individual was predicted as having low-literacy. A summary of the predictive accuracy of this measure is given in the Table 2. Based on the results in Table 2, the sensitivity of the test is 75.0 percent with 95 percent confidence intervals of 52.4 to 93.6 percent. The specificity of the test is 82.9 percent with 95 percent confidence intervals of 73.0 to 90.3 percent.

The analysis for the Computer Skill Score is shown in Table 3. The only significant variable in predicting the Computer Skills Score was how comfortable a person was with the computer mouse.

Model for Predicting Reading Literacy. In order to customize the MADELINE system to support the

Table 1. Stepwise Multiple Regression Analysis for Sight Word Score with a Linear Educational Level. $R^2 = 0.5492$.

Source	DF	Sum of Squares	Mean Square	F	Prob>F
Regres.	5	19656.249	3931.3	22.90	0.0001
Error	94	16136.741	171.7		
Total	99	35623.702			

Variable	Parameter Estimate	Standard Error	F	Prob>F
Age	-0.42952	0.10677	16.18	0.0001
Education level	4.47234	1.29510	11.92	0.0008
# of books read	3.52855	1.10437	10.21	0.0019
Low education	-9.78919	4.91166	3.97	0.0492
Reading preference	-7.58519	3.12673	5.89	0.0172

Table 2. Predictive Accuracy of the Linear Regression Model

Predicted Literacy Level from Model	Actual Literacy Level (Sight Word Test)		
		Low	High
	Low	14	14
	High	4	68

Table 3. Stepwise Multiple Regression Analysis for Computer Skill. $R^2 = 0.5701$.

Source	DF	Sum of Squares	Mean Square	F	Prob>F
Regres.	1	24521.860	24521.9	107.42	0.0001
Error	81	18490.935	228.3		
Total	82	43012.795			

Variable	Parameter Estimate	Standard Error	F	Prob>F
Mouse use	21.25300	2.05060	107.42	0.0001

reading literacy of each individual, we created a regression model to generate a literacy score based on the following questions from the exploratory study:

- Initial score based on last grade of school completed:

Answer	Score
a) Less than high school (0-8)	72.87
b) Some HS but did not graduate	84.66
c) High school graduate or GED	89.02
d) Tech. school education	93.37
e) Some college but no degree	93.37
f) College graduate	97.73
g) Advanced degree	97.73
- Correction for age: subtract ($\text{age} \times .412$)
- Correction for "How often do you read books":

a) never	add 0.00
b) less than 1/week	add 3.45
c) 1/week	add 6.91
d) a few times a week	add 10.36
e) every day	add 13.81
- Correction for reading preference:

a) I'd like to read them myself	subtract 0.00
b) I'd like the questions read to me	subtract 7.44

If the total score is less than 75, then the individual is assigned to the low-literacy category.

For computer literacy, we found that the question regarding a subject's comfort with the use of a

computer mouse was most predictive of computer skill. Subjects that responded that they were not comfortable with a mouse were categorized as having low computer skill. Those who responded that they were very comfortable were classified as having high computer skill. Subjects who responded that they were "somewhat comfortable" with a computer mouse were asked a follow-up preference question regarding whether or not they wanted to use a touch screen or a computer mouse and a keyboard to enter their information. Subjects selecting the former answer were classified as having low computer skill and subjects selecting the latter answer were classified as having high computer skill.

Use of the Assessment Tool. The assessment tool was used to categorize by reading literacy and computer skill the 705 subjects who participated in the MADELINE system evaluation study. In this context, the subjects were assigned to one of four levels of support using the assessment tool: low reading, low computer; high reading, low computer; low reading, high computer; and high reading, high computer. By using the support levels assigned by the assessment tool, the question completion rate of subjects with low reading literacy was made comparable to the completion rate of subjects with high reading literacy. In addition, both low and high reading literacy subjects completed significantly more questions using MADELINE relative to a paper form.

Model Validation. The results using the WRAT3 absolute score (a continuous variable without regard to grade or age) showed that the model distinguished between high and low literacy subjects with a high level of statistical significance ($p=0.0025$). Table 3 gives the mean, median, inter-quartile range, and standard deviation of the WRAT3 Absolute score. The scores are summarized by literacy group.

An absolute WRAT3 score of 506 corresponds to the target reading level of our low literacy content, i.e., fifth grade or lower. Thus, with low literacy set as both a WRAT3 score less than 506 and a model-generated reading literacy score of less than 75, the performance of our tool corresponds to an estimated area-under-the-curve of 0.63 with a statistically significant 95 percent confidence interval of 0.53 to 0.72. Receiver operating characteristic curve analysis was not applicable since both parameters for reading

Table 3. WRAT3 Absolute Scores by Literacy Group for Subjects Participating in the MADELINE Evaluation.

Metric	Low Literacy Group	High Literacy Group
Mean	508.4	520.2
Median	509*	522*
Inter-quartile range	502, 515	509, 530
Standard deviation	12.7	13.7

*p-Value for Difference = 0.0025

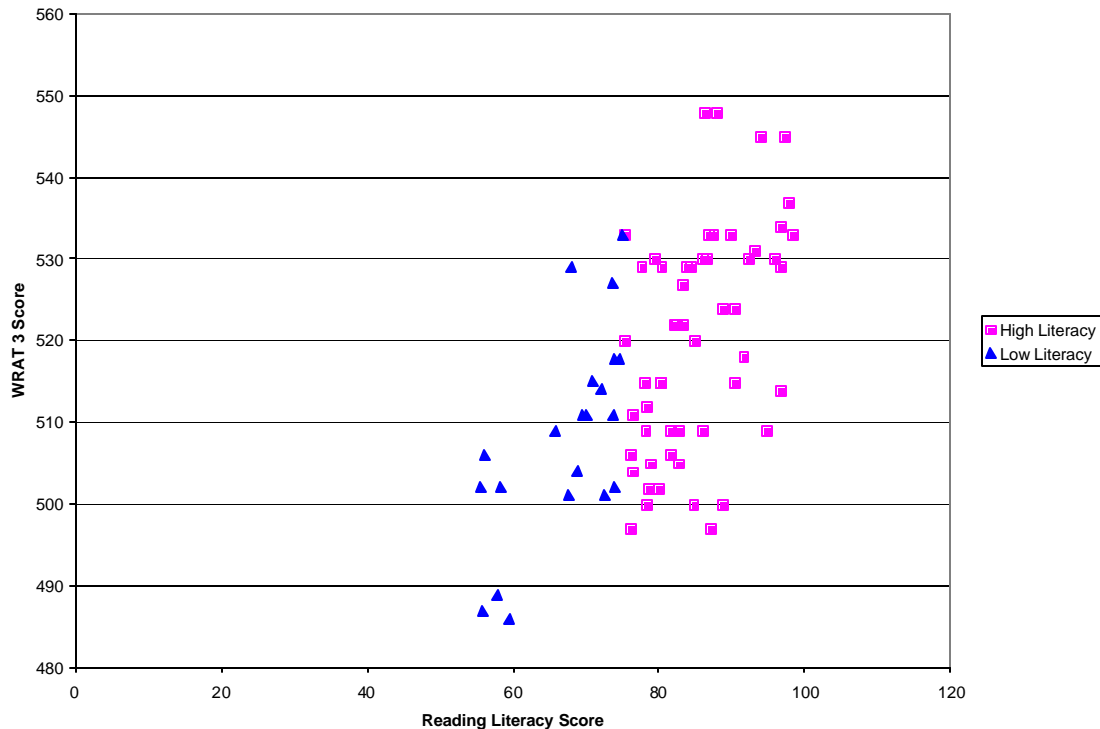


Figure 2. WRAT3 Absolute Score vs. Reading Literacy Score Calculated from Model

literacy were preset and could not be adjusted to further optimize the tool's performance.²⁴ A scatter plot of the reading literacy score generated by the model vs. the WRAT3 score is depicted in Figure 2.

DISCUSSION

The results of this study demonstrate that the skill assessment tool developed for the MADELINE project is able to categorize patients with regard to high or low reading literacy level and high or low computer skill with a moderate degree of accuracy. The effectiveness of the tool was best demonstrated when used in the MADELINE application, where it enabled low literacy users to enter clinical information on a par with high literacy patients. The advantage of our assessment tool is that it can be administered through a computer in less than two minutes without requiring any special training or expertise. Thus, this tool can be very useful to make a rapid determination of skill, so that patients can receive the appropriate data collection interface as was done in this project. This tool could also be employed in other settings in which a rapid categorization of patients based on reading ability or skill with computers would be necessary.

When compared with tests for reading literacy, the tool was able to discriminate between high and

low literacy patients with statistical significance. While this performance was not stellar, we expected our literacy assessment tool to perform less well when applied to the confirmatory data set because the tool was optimized for the original exploratory data set. In addition, discordance between the assessment tool and the WRAT3 may in part reflect the fact that the skills addressed by the assessment tool may not correlate directly with the literacy assessment of the WRAT3. In other words, even though the WRAT3 is a validated literacy assessment tool, it may not be the ideal "gold standard" for assessing the skills needed to enter clinical information into a computer. The ability of the assessment tool to enable the groups it assigned to low and high literacy to complete questions at an equivalent level shows that the tool was effective for identifying which users needed greater support for reading literacy. In contrast, the ability of the high and low literacy groups to contribute information was significantly different when paper forms were used.

CONCLUSION

The results of this study demonstrate that it is feasible to quickly assess, with reasonable accuracy, a patient's reading literacy and computer skill in order to appropriately adapt a computer-administered

patient interview. This method could be applied to a health maintenance and risk assessment questionnaire or any of a variety of questionnaires related to the provision of health care services. Application of these methods have the potential to improve the quality of care provided to patients, particularly those with low literacy, because they will be empowered to provide more comprehensive information about their health status to their physicians.

Acknowledgements

The authors wish to thank Marci K. Campbell, PhD, of the University of North Carolina (Chapel Hill, NC) for contributing to the design and evaluation of the assessment tool, and to Jennifer M. Arbanas, MS, and Nichole Berke, MS, of Duke University for conducting the assessment tool evaluation studies. We are also grateful to Susan C. Lowell, MA, of Learning and Development Associates (Morrisville, NC) for guidance regarding literacy assessment. This work was funded in part by grant R18 HS09706 from the Agency for Healthcare Research and Quality.

References

- Slack WV, Hicks GP, Reed CE, Van Cura LJ. A computer-based medical-history system. *N Engl J Med* 1966;274:194-198.
- Mayne JG, Martin MJ. Computer-aided history acquisition. *Med Clin North Am* 1970;54:825-33.
- Grossman JH, Barnett GO, McGuire MT, Swedlow DB. Evaluation of computer-acquired patient histories. *JAMA* 1971;215:1286-1291.
- Stead WW, Heyman A, Thompson HK, Hammond WE. Computer-assisted interview of patients with functional headache. *Arch Intern Med* 1972;129:950-955.
- Card WI, Nicholson M, Crean GP, et al. A comparison of doctor and computer interrogation of patients. *Int J Biomed Comput* 1974;5:175-87.
- Collen MF. Patient data acquisition. *Medical Instrumentation* 1978;12:222-225.
- Quaak MJ, Westerman RF, van Bommel JH. Comparisons between written and computerised patient histories. *BMJ Clinical Research Ed* 1987;295:184-190.
- Slack WV, Safran C, Kowaloff HB, Pearce J, Delbanco TL. Be well: a computer-based health care interview for hospital personnel. *Proc Annu Symp Comput Appl Med Care* 1993:12-16.
- Bachman JW. The patient-computer interview: a neglected tool that can aid the clinician. *Mayo Clin Proc* 2003;78:67-78.
- Sutherland LA, Campbell M, Ornstein K, Lobach DF. Development of an adaptive multimedia program to collect patient health data. *Am J Prev Med* 2001;21:320-324.
- Lobach DF, Arbanas JM, Mishra DD, Wildemuth B, Campbell M. Collection of information directly from patients through an adaptive human-computer interface. *Proc AMIA Symp* 1999:1087.
- Yarnold PR, Stewart MJ, Stille FC, Martin GJ. Assessing functional status of elderly adults via microcomputer. *Percept Mot Skills* 1996;82:689-90.
- Pearson J, Jones R, Cawsey A, et al. The accessibility of information systems for patients: use of touchscreen information systems by 345 patients with cancer in Scotland. *Proc AMIA Symp* 1999:594-8.
- Wofford JL, Currin D, Michielutte R, Wooford MM. The multimedia computer for low literacy patient education: a pilot project of cancer risk perceptions. *Med Gen Med* 2001;3:23.
- Snelbaker AJ, Wilkinson GS, Robertson GJ, Glutting JJ. Wide Range Achievement Test 3 (WRAT 3). In: Dorfman WI, Hersen M, editors. *Understanding psychological assessment: perspectives on individual differences*. New York: Kluwer Academic/Plenum Publishers; 2001. p. 259-274.
- Torgesen JK, Wagner RK, Rashotte CA. *Test of word reading efficiency*. Austin (TX): PRO-ED, Inc.; 1999.
- TenHave TR, Van Horn B, Kumanyika S, Askov E, Matthews Y, Adams-Campbell LL. Literacy assessment in a cardiovascular nutrition education setting. *Patient Educ & Couns* 1997;31:139-50.
- Nath CR, Sylvester ST, Yasek V, Gunel E. Development and validation of a literacy assessment tool for persons with diabetes... including commentary by Kaplan SM. *Diabetes Educator* 2001;27:857-64.
- Balajthy E. Information technology and literacy assessment. *Reading & Writing Q: Overcoming Learning Difficulties* 2002;18:369-73.
- National Center for Education Statistics (U.S. Department of Education). *National Adult Literacy Survey* 1992.
- Torkzadeh G, Koufteros X. Factorial validity of a computer self-efficacy scale and the impact of computer training. *Educ & Psych Meas* 1994;54:813-821.
- Murphy CA, Coover D, Owen SV. Development and validation of the Computer Self-Efficacy Scale. *Educ & Psych Meas* 1989;49:893-899.
- Rosner B. *Fundamentals of biostatistics*. 4th ed. Belmont (CA): Duxbury Press; 1995. p.558-562.
- Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 1982;143:29-36.